Forensic DNA Typing

Technology Focus Day on Human Identification by DNA Fingerprinting
In-Q-Tel
January 22nd, 2008
Arlington, VA

Dr. Peter M. Vallone, Biochemical Science Division,
National Institute of Standards and Technology
Gaithersburg, Maryland

Outline
• NIST
• Basics of DNA Typing
• Paternity Testing
  • Y-STRs
  • miniSTRs
• SNPs
• Examples

NIST History and Mission
• National Institute of Standards and Technology (NIST) was created in 1901 as the National Bureau of Standards (NBS). The name was changed to NIST in 1988.
• NIST is part of the U.S. Department of Commerce with a mission to develop and promote measurement, standards, and technology to enhance productivity, facilitate trade, and improve the quality of life.
• NIST supplies over 1,300 Standard Reference Materials (SRMs) for industry, academia, and government use in calibration of measurements.
• NIST defines time for the U.S.

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Location of NIST (Gaithersburg, MD)

Location of NIST

Washington, D.C.

BWI Airport

Richmond, VA

I-95

I-66

~30 miles

Dulles Airport

Reagan National Airport

FBI Lab

NIST

National Institute of Justice

The Research, Development, and Evaluation Agency of the U.S. Department of Justice

Current Areas of Effort with Forensic DNA

• Standards
  – Standard Reference Materials
  – Standard Information Resources (STRBase website)
  – Interlaboratory Studies

• Technology
  – Research programs in SNPs, miniSTRs, Y-STRs, mtDNA, qPCR
  – Assay and software development, expert system review

• Training Materials
  – Review articles and workshops on STRs, CE, validation
  – PowerPoint and pdf files available for download

http://www.cstl.nist.gov/biotech/strbase/NIJprojects.htm

STRBase
Short Tandem Repeat DNA
Internet Database

Recent Additions
- Forensic SNP Information (will be official site for ISFG SNP information)
- NIST publications and presentations as pdf files

We Regularly Update
- Reference List
- Variant Alleles
- Addresses for Scientists
- Links to Other Web Sites
- Y-STR Information

We will continue to add downloadable PowerPoint files that can be used for training purposes

http://www.cstl.nist.gov/biotech/strbase

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Basics of Forensic DNA Testing

Methods for Human Identification

Fingerprints have been used since 1901

DNA since 1986

General Characteristics of Genomic DNA

- Each person has a unique DNA profile (except identical twins)
- Each person's DNA is the same in every cell (DNA from skin cells will match DNA from blood cells)
- An individual's DNA profile remains the same throughout life
- Half of your DNA comes from your mother and half from your father
Forensic DNA Testing

Probe subsets of genetic variation in order to differentiate between individuals.

DNA typing must be done efficiently and reproducibly (information must hold up in court)

Typically, we are not looking at genes – little/no information about race, predisposition to disease, or phenotypical information (eye color, height, hair color) is obtained

Applications of Human Identity Testing

- Forensic cases -- matching suspect with evidence
- Paternity testing -- identifying father
- Missing persons investigations
- Military DNA “dog tag”
- Convicted felon DNA databases
- Mass disasters -- putting pieces back together
- Historical investigations and genetic genealogy

Involves generation of DNA profiles usually with the same genetic markers and then MATCHING TO REFERENCE SAMPLE

DNA Testing Requires a Reference Sample

A DNA profile by itself is fairly useless because it has no context…

DNA analysis for identity only works by comparison – you need a reference sample

Crime Scene Evidence compared to Suspect(s) (Forensic Case)
Child compared to Alleged Father (Paternity Case)
Victim’s Remains compared to Biological Relative (Mass Disaster ID)
Soldier’s Remains compared to Direct Reference Sample (Armed Forces ID)
Steps in DNA Analysis

**Steps Involved**
- Collection
- Specimen Storage
- Extraction
- Quantitation
- Multiplex PCR
- STR Typing
- Interpretation of Results
- Database Storage & Searching
- Calculation of Match Probability

**Steps in DNA Analysis**
Usually 1-2 day process (a minimum of ~5 hours)

1. **Sample Collection & Storage**
2. **DNA Extraction**
3. **Multiplex PCR Amplification**
4. **DNA Quantitation**
5. **Slot Blot**
   - 1 ng
   - 0.3 ng
   - 1 ng
   - 1 ng
   - 0.7 ng
   - 0.5 ng
   - 0.5 ng
   - No DNA

**If a match occurs**, comparison of DNA profile to population allele frequencies to generate a case report with probability of a random match to an unrelated individual

**STR Typing**
DNA separation and sizing

**Technology**
- Biology
- Genetics
- DNA Database Search
- DNA Database Search

**Unfortunately, current DNA testing cannot be performed as quickly as a commercial break...**

The instruments on CSI are real – they just do not collect data as quickly as shown on TV

**Sources of Biological Evidence**
- Blood
- Semen
- Saliva
- Urine
- Hair
- Teeth
- Bone
- Tissue

**Blood Sample**
Only a very small amount of blood is needed to recover DNA

Best results with >100 cells, but DNA profiles can be recovered from as little as a single cell
DNA in the Cell

- Target Region for PCR
- Double stranded DNA molecule
- 22 pairs + XX or XY
- ~3 billion total base pairs

Human Genome

- 23 Pairs of Chromosomes + mtDNA
- Located in cell nucleus
- Located in mitochondria (multiple copies in cell cytoplasm)
- Nuclear DNA: 3.2 billion bp
- mtDNA: 16,569 bp

What Type of Genetic Variation?

- Length Variation: short tandem repeats (STRs)
  - CTAGTCGT(GATA)(GATA)(GATA)GCGATCGT
- Sequence Variation: single nucleotide polymorphisms (SNPs)
  - insertions/deletions
  - GCTAGTCGATGCTCGTGATAGC

Short Tandem Repeat (STR) Markers

An accordion-like DNA sequence that occurs between genes

TCCCAAGCTCTTCTCCCTATCAGACAAGAGACA
GCTGATAGATAGATAGATAGATAGATAGATAGATAGATAG
TAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGAT
ACATGCTAGATAGATAG

= 12 GATA repeats ("12" is all that is reported)

The number of consecutive repeat units can vary between people

Target region (short tandem repeat)

The FBI has selected 13 core STR loci that must be run in all DNA tests in order to provide a common currency with DNA profiles

13 CODIS Core STR Loci with Chromosomal Positions

What is a DNA Profile?

Unique regions of the human genome are targeted

These regions consist of a few hundred base pairs

The regions are copied by the polymerase chain reaction (PCR) – billions of exact copies

The copied fragments now contain fluorescent dyes for detection

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
PCR Process

Separate strands (denature)

Add primers (anneal)

Make copies (extend primers)

Repeat Cycle, Copying DNA Exponentially

Forward Primer

Reverse Primer

Starting DNA Template 80-500 bases

PCR Copies DNA Exponentially through Multiple Thermal Cycles

In 30 cycles at 100% efficiency, 1.07 billion copies of targeted DNA region are created

What is a DNA Profile?

The labeled fragments are separated (based on size) and detected on a gel or capillary electrophoresis instrument

~2 hours or less

Peak size ranges from 100 - 350 base pairs

Peaks represent labeled DNA fragments separated by electrophoresis. This ‘profile of peaks’ is unique for an individual – a DNA type
Capillary Electrophoresis (CE)

Fill with Polymer Solution

Outlet (anode)

Argon Ion Laser

50-100 µm x 27 cm

Burn capillary window

Inlet (cathode)

- DNA Separation occurs in minutes...

5-20 kV

Data Acquisition and Analysis

Sample tray moves automatically beneath the cathode end of the capillary to deliver each sample in succession

DNA Separation occurs in minutes...

Sample tray moves automatically beneath the cathode end of the capillary to deliver each sample in succession

Capillary Electrophoresis Instrumentation

AB 310
single capillary

AB 3130
16-capillary array

Newer Instruments
• 3730 – 48 capillaries
• 3730xl – 96 capillaries
• Chip platforms

AB = Applied Biosystems (Foster City, CA)

Identifiler STR Kit

Information is tied together with multiplex PCR and data analysis

The Random Match Probability (RMP) is over 1 in 800 trillion for unrelated individuals

The chance of someone else having this exact same profile
Product Rule

For heterozygous loci

\[ P = 2pq \]

\( P \) = probability; \( p \) and \( q \) are frequencies of allele in a given population

Example: For the locus D3S1358 an individual is 15,18 with frequencies of 0.2825 and 0.1450 respectively

\[ P = 2(0.2825)(0.1450) = 0.0819 \text{ or 1 in 12} \]

For 5 loci the Profile Probability = \( (P_1)(P_2) \ldots (P_n) \)

\[ = (0.0819)(0.0875)(0.0687)(0.0245)(0.0984) \]

\[ = 0.000001187 \text{ or 1 in 842,539} \]

Companies Supply Allelic Ladders in STR Kits to Aid Interlaboratory Consistency

Profiler Plus kit allelic ladders (Applied Biosystems)

Allelic Ladders

PCR Product Size (bp)

Sample #1

Sample #2

Companies Supply Allelic Ladders in STR Kits to Aid Interlaboratory Consistency

Profiler Plus kit allelic ladders (Applied Biosystems)
Data Format

<table>
<thead>
<tr>
<th>AMEL</th>
<th>CSF1PO</th>
<th>FGA</th>
<th>TH01</th>
<th>TPOX</th>
<th>VWA</th>
<th>D3S1358</th>
<th>D5S818</th>
</tr>
</thead>
<tbody>
<tr>
<td>X,Y</td>
<td>11,12</td>
<td>19,21</td>
<td>6,7</td>
<td>9,8</td>
<td>15,18</td>
<td>14,16</td>
<td>10,13</td>
</tr>
</tbody>
</table>

The profiles are reviewed by analysts

The number of repeats observed for each locus is tabulated

This data format is stored in databases and used for comparisons/matches

A Brief History of Forensic DNA Typing


1996: STR typing with CE is routinely used worldwide

US National Database launched (October 13, 1998)

1998: STR = short tandem repeat

2000: Y-chromosome STRs

2002: STR multiplex kits

2004: First commercial fluorescent

2008: First miniSTR Commercial kit

2008: DNA is an important part of the criminal justice system

STR = short tandem repeat

Paternity Testing
Paternity Testing

- The use of DNA testing methods for determining paternity also relates to:
  - Mass disasters
  - Missing persons investigations
  - Familial matching
  - Genetic Genealogy

Our DNA Comes from our Parents

- Father's Sperm
- Mother's Egg
- Child's Cell

Different Inheritance Patterns

- CODIS STR Loci
- Lineage Markers
  - Autosomal (passed on in part, from all ancestors)
  - Y-Chromosome (passed on complete, but only by sons)
  - Mitochondrial (passed on complete, but only by daughters)
Autosomal Paternity Example

- Dad
- Me
- Sister

Y-STRs

- Similar to autosomal STRs just located on the Y-Chromosome
- Since only males possess a Y-Chromosome, these markers are useful in male-female mixtures (sexual assault cases)
- A limitation of the Y-STRs lies in that do not have the discrimination capacity of autosomal STRs (no recombination)
Traits found on the Y-Chromosome

- Channel Flipping (FLP)
- Joke Telling (GOT-1)
- Air Guitar (RIF)
- Spitting (P2E)
- Selective hearing loss (HUH)

Y-Chromosome Paternity Example

- Maternal Uncle
- Uncle
- Dad
- Me

Y-Chromosome Case Example

- Can be excluded
- Suspect 1
- Can not be excluded
- Suspect 2
- Can not be excluded
- Suspect 3
- Excluded
Modern Use of Y-STR Testing
Captured December 13, 2003

Matching Y-STR Haplotype Used to Confirm Identity
( along with allele sharing from autosomal STRs)

Is this man really Sadaam Hussein?
Killed July 22, 2003

Uday and Qusay Hussein

Sadaam Hussein

Box 23.1, p. 534

Historical Investigation DNA Study
(Matching Relatives to Remains or Relatives to Relatives)

Thomas Jefferson II

Field Jefferson (Uncle)

Paul Jefferson

President Thomas Jefferson

Eston Hemings

Jefferson Y Haplotype

Jefferson Y Haplotype

Same Y Haplotype

Different Y Haplotype


Butler, J.M. (2005) Forensic DNA Typing, Figure 17.4, Academic Press

miniSTRs
miniSTRs

- Simply a smaller PCR product size
- Typically less than ~200 base pairs
- Contains the same information as a traditional STR (repeat length)
- Useful for typing degraded DNA samples
- New loci helpful for missing persons paternity testing/mass disasters

DNA Degradation

A MiniFiler Kit Profile

Commercial miniSTR test from AB focusing on reducing PCR products length for 8 of the larger Identifiler loci

Initial work developing miniSTRs was started for WTC identifications

John Butler (NIST)
Bruce McCord (FIU)
Bode Technology Group

Data from Ms. Becky Hill (NIST)
World Trade Center – Phase I Summary

12,392 Bone samples processed
Over 6800 profiles
3,405 Full profiles (13 STR loci)
2,143 High partial profiles (>7 STR loci)
2,670 Low partial profiles (<7 STR loci)
4,174 No loci

Over 6800 profiles miniSTRs are helping here

Final 20% of WTC victims identified were based on a miniSTR technique pioneered at NIST

SNPs

What Type of Genetic Variation?

• Length Variation
  short tandem repeats (STRs)
  CTAGTCGT(GATA)(GATA)(GATA)GCGATCGT

• Sequence Variation
  single nucleotide polymorphisms (SNPs)
  insertions/deletions
  GCTAGTCGATGTC(G/A)GCGATGCTGAGC
Potential Use of SNPs in Forensic DNA Testing

- Human Identification (need 50+ loci)
- Predicting Geographical Origin
- Prediction of Phenotypical Information – Hair color, eye color etc
- Evolutionary studies
- May be cheaper/faster
- Can be run on higher throughput platforms
- Current databases are for the 13 CODIS loci

Rapid PCR

- Existing commercial STR typing kits are not optimized for rapid PCR
- Challenge for miniaturize STR typing platforms – since they are tied into the commercial kits/loci
- Fewer loci and smaller amplicon size favor rapid multiplex PCR
- We have well characterized miniSTR panels

Collaborations with:
Dr. Michael Gaitan (NIST) – microwave thermal heating
Dr. Eugene Tan (Network Biosystems) – chip platforms
Rapid Thermal Cycling

- Evaluate faster polymerases
- Test with miniSTRs
  - primer concentrations can be adjusted and PCR primer sequences are known
- Use standard cycler (GeneAmp 9700), tubes
- Examine shortened dwell times and adenylation soak
- Study limitations in terms of PCR amplification speed when examining multiplex STR assays

Rapid PCR 28 cycles in 36 minutes on a AB 9700

NYC Prize

- In the months ahead, we will also challenge the private sector to speed up DNA fingerprinting so that when DNA is left behind, officers can identify suspects more quickly and avoid wrongful arrests. And to do this, we will establish a six-figure prize for anyone who can invent a device tailored to the NYPD that analyzes DNA right at the crime scene. It's just one more way we are trying to bring private sector innovation into the public sector
- The City will lobby Albany to require DNA fingerprinting from everyone who is arrested
Forensic DNA in the News

DNA EXONERATIONS BY YEAR IN THE U.S.

The Genographic Project

- Different populations carry distinct markers. Following them through the generations reveals a genetic tree on which today's many diverse branches may be followed ever backward to their common African root.
- Our genes allow us to chart the ancient human migrations from Africa across the continents.
- Funded $50 million for 5 years by IBM and National Geographic.
- Will gather and run DNA samples from ~100,000 people around the world with Y-SNPs and mtDNA.
Tsunami Survivor “Baby 81” Connected to His Parents with DNA

NEW YORK (AP) — The parents of the infant tsunami survivor nicknamed "Baby 81" say they found it difficult to feel overjoyed about their reunion in the midst of so much tragedy.

The 4-month-old Sri Lankan baby and his parents, who were reunited after court-ordered DNA tests proved their relationship, appeared on ABC's "Good Morning America" Wednesday, a day after their 20-hour-long flight landed in New York.

Identification of Remains from Former Yugoslavia

>90,000 family reference samples collected
>17,000 bones identified as of April 2007

DNA testing is performed on 100s of bones collected each week from mass graves in Bosnia and Croatia to help in the re-association of remains

The New York Times

April 11, 2006

Hurricane Katrina Victim Identification Being Performed with DNA

In Attics and Rubble, More Bodies and Questions

August Blanchard and an aunt, Shirley, looking into the living room where the decomposed remains of his mother, Charlene, were found in February on Reynes Street in the Lower Ninth Ward of New Orleans.


DNA Testing Assists Immigration

DNA Tests Offer Immigrants Hope or Despair

Mary K. Mount, a DNA testing expert for the A.A.B.B.—formerly known as the American Association of Blood Banks—estimates that about 75,000 of the 390,000 DNA cases that involved families in 2004 were immigration cases. Of those, she estimates, 15 percent to 20 percent do not produce a match.

Negative results can suggest an effort to bring in illegal immigrants or distant relatives, officials say, though they note that requests for DNA tests deter illicit activities.


Sandra and Balfour Francis of Brooklyn, with a photograph of Nickiesha, who is in Jamaica. Last year, DNA tests showed she is not his daughter.

The New York Times

April 10, 2007

Lawyers for Duke Players Say DNA Clears Team


Armed Forces DNA Repository

>4.5 million bloodstain cards on file from members of U.S. military

Are being used to identify remains from combat casualties

Located in Gaithersburg, Maryland
Tomb of the Unknown Soldier

- Armed Forces DNA Identification Laboratory (AFDIL) (Rockville, MD)
- In June 1998 AFDIL identified Michael J. Blassie as the Vietnam Unknown in the Tomb of the Unknown Soldier (located in Arlington National Cemetery)
- There will be no more “unknown” soldiers.

August 17, 1998
FBI Report on Analysis of Stain on Monica Lewinsky’s Blue Dress

Results Announced in Anna-Nicole Smith DNA Parentage Test

Larry Birkhead receives a handshake from DNA expert Dr. Michael Baird outside the court after a paternity hearing in Nassau, Tuesday, April 10, 2007.
Genetic Genealogy

Stalking Strangers' DNA to Fill in the Family Tree

NIST Human Identity Project Team

John Butler (Project Leader) Margaret Kline Pete Vallone

Dave Duewer Jan Redman Amy Decker Becky Hill

Funding: Interagency Agreement 2003-JJ-R-029 between National Institute of Justice (NIJ) and NIST Office of Law Enforcement Standards (OLES)